

DOCUMENT RESUME

ED 170 331

TH 008 557

AUTHOR Misanchuk, Earl R.
TITLE Relationship Among the Number of Sub-Tests; Skewness, Kurtosis, and Size of Population; And Magnitude of Errors of Estimate in Multiple Matrix Sampling. (Revised Version).
PUB DATE Mar 78
NOTE 30p.; Paper presented at the Annual Meeting of the American Educational Research Association (62nd, Toronto, Ontario, Canada, March 27-31, 1978)
EDRS PRICE MF01/PC02 Plus Postage.
DESCRIPTORS Adults; Analysis of Variance; *Item Sampling; Participant Characteristics; Personality Assessment; Probability; Research Design; Scores; *Standard Error of Measurement; Statistical Analysis; Statistical Studies
IDENTIFIERS California Psychological Inventory; Test Length

ABSTRACT

Multiple matrix sampling of three subscales of the California Psychological Inventory was used to investigate the effects of four variables on error estimates of the mean (EEM) and variance (EEV). The four variables were examinee population size (600, 450, 300, 150, 100, and 75); number of subtests, (2, 3, 4, 5, 6, and 7), hence the number of items per examinee; skewness; and kurtosis. The relationship of population size, number of subtests, and skewness and kurtosis of each subscale/population combination to the EEM and EEV was examined, by performing an analysis of variance on stepwise multiple linear regression. A total of 324 individual multiple matrix samples were involved. EEM was not significantly affected by any variable studied; but EEV, in contrast to previous studies, was significantly affected by skewness and kurtosis. (CP)

 * Reproductions supplied by EDRS are the best that can be made *
 * from the original document. *

ED170331

U.S. DEPARTMENT OF HEALTH,
EDUCATION & WELFARE
NATIONAL INSTITUTE OF
EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY.

RELATIONSHIP AMONG THE NUMBER OF SUB-TESTS; SKEWNESS, KURTOSIS,
AND SIZE OF POPULATION; AND MAGNITUDE OF ERRORS OF ESTIMATE
IN MULTIPLE MATRIX SAMPLING

Earl R. Misanchuk¹

The University of Saskatchewan

"PERMISSION TO REPRODUCE THIS
MATERIAL HAS BEEN GRANTED BY

Carl R. Misanchuk

TO THE EDUCATIONAL RESOURCES
INFORMATION CENTER (ERIC) AND
USERS OF THE ERIC SYSTEM."

Paper presented at the annual meeting of the
American Educational Research Association,
Toronto, March, 1978.

(REVISED VERSION)

SESSION 26.13

TM008 557

Abstract

Post mortem multiple matrix sampling of three subscales of a personality test was used to evaluate the effects of population size, the number of sub-tests used, and population skewness and kurtosis on errors of estimate of the mean (EEM) and variance (EEV). EEM was not significantly affected by any variable studied, but EEV was significantly affected by population skewness and the interaction of skewness and kurtosis.

A number of studies have demonstrated that multiple matrix sampling (MMS) provides more precise estimates of population means and variances than conventional examinee sampling (Barcikowski, 1972; Cook & Stufflebeam, 1967; Lord, 1962; Plumlee, 1972; Pugh, 1971; Shoemaker, 1970a, 1970b, 1973), and a growing number of studies demonstrate the influence on errors of estimate of various data base characteristics and sampling plans (e.g., Barcikowski, 1974; Moy & Barcikowski, 1974; Shoemaker, 1970b, 1971, 1972, 1973). Typically, such demonstrations are done through post mortem sampling, in which an extant data base is sampled, either with MMS or examinee sampling, and estimates of parameters are generated. The estimates are then compared to the (known) parameters to determine the magnitude of error associated with the sampling scheme.

Equations developed for MMS by Lord and Novick (1968) are based on infinitely large population and sub-test sizes, but Shoemaker (1973) presents refinements designed to allow MMS to be used with a wide variety of sampling plans. Shoemaker implies that his equations can be used with substantially smaller numbers of examinees per sub-test than those originally given by Lord and Novick (1973, p. 34), but no explicit guidelines are given. Of the few guidelines that do exist with respect to the design of sampling plans, most are, at this juncture, general enough to be of limited use in making decisions when designing sampling schemes. Indeed, some of the design prescriptions are conflicting, in the light of finite item and examinee pools, viz., "always have at least 6 to 8 sub-tests," but "have more items per sub-test and fewer sub-tests [rather] than many sub-tests with few items each (Shoemaker, 1973, p. 22)." Barcikowski (1974) supports the recommendation to have a large number of sub-tests

when estimating the mean of a population, and but contradicts Shoemaker when he recommends the use of small sized sub-tests.

The Number of Sub-tests

Given a finite item pool, the number of items per sub-test can be varied in two ways: by changing the number of sub-tests, or by putting some items into more than one sub-test. (The theoretical framework of MMS does in fact permit the second alternative but such an action would only seem warranted in the light of evidence that a particular number of sub-tests represents an optimal sampling plan.) A number of studies have examined the effects of different sampling schemes. They typically manipulate the number of sub-tests, the number of items per sub-test, and/or the number of examinees per sub-test to generate estimates of parameters (e.g., Barcikowski, 1974; Moy & Barcikowski, 1974; Shoemaker, 1970a, 1970b, 1971, 1972), but none has set out to evaluate the effect on errors of estimates of systematically varying the number of sub-tests across several data bases (distributions). Perhaps design guidelines with respect to the optimum number of sub-tests to be used could be made more directive through extended experience with the MMS procedure being applied to different sampling configurations.

Population Size

Although the procedure of MMS has been used with small populations (N) (e.g., see Shoemaker & Okada, 1970), the demonstrations of superiority of

precision of MMS have usually involved relatively large N 's (e.g., see Barcikowski, 1972; Cook & Stufflebeam, 1967; Lord, 1962; Scheetz & Forsyth, Note 1; Shoemaker, 1970a, 1970b). Moreover, those studies have not deliberately set out to determine the relative efficacy of MMS on various population sizes. Typically, the sampling studies compare the estimates generated through varying sampling plans (which influences the effective size of N , the population size, by varying the sizes of t , the number of sub-tests, and n , the number of examinees per sub-test) to the normative distribution (with a constant maximal value of N). That is, as the size of an effective population (defined as the product tn) varies, so do the magnitudes of its parameters, which point seems to have been ignored in the studies examining the effects of various sampling schemes.

There are many situations in which MMS might be applied where the researcher has no control over the size of the population (e.g., as in collecting survey data), and the possibility exists that various population sizes could be optimally sampled through different sampling schemes. Notwithstanding the foregoing, MMS could be applied to small groups with greater confidence if there were empirical evidence of its ability to maintain the precision of its estimates with considerably smaller populations. Thus the efficacy of MMS with various population sizes seems worthy of study.

Skewness

Another population characteristic whose effect on errors of estimates under

MMS has been examined is skewness. Shoemaker (1973) relates the effectiveness of various sampling plans in reducing standard errors of estimate to the skewness of the population; however, Forsyth (1976) takes exception to Shoemaker's claims and offers theoretical evidence against them. Instead, Forsyth implies, the conclusions may be peculiar to Shoemaker's definition of skewness and to the data base used.

One limiting, but understandable, characteristic typifying MMS studies to date has been their emphasis on within-data-base replications. It is not unusual to find sampling schemes replicated 100 or even 300 times. While this design element improves the precision of the estimates (i.e., it minimizes the error in estimating the errors of estimate), it usually sacrifices generalizability by confining the examination to a selected data base (Shoemaker, 1970b is an exception--he used two data bases, but did not make direct comparisons based on the differentiation). Using an analysis-of-variance analogy, studies to date have emphasized lowering the within-cells variance at the expense of examining the between-groups variance. A study that replicated many times within a data base but also used many data bases would permit much more confident generalizations about the effects of certain population characteristics on errors of estimates. The major difficulty in performing such a study is pragmatic--even with present-day computers, it is probably computationally and economically practical only to address one or the other of the two conflicting concerns in more than a limited way in a given study.

The present study attempts to break away from the traditional design of

MMS studies to examine errors of estimates as affected by various population characteristics in the context of several subscales. In doing so, for reasons of economy, it admittedly sacrifices some of the precision inherent in the more traditional approach.

Method

The present study employed post mortem sampling on a data base consisting of 600 sets of responses to the California Personality Inventory (CPI) that had been administered to business executives as part of an on-going selection and placement enterprise (see Scissons, 1977). In an attempt to expand generalizability, three subscales of the CPI, the Dominance (Do), Self-Acceptance (Sa), and the Communalinity (Cm), were selected arbitrarily for use in the study.

The 600 cases were randomly ordered, and six populations of $N = 600, 450, 300, 150, 100,$ and 75 testees, respectively, were created for each subscale by selecting the first N cases from the list. The entire process was repeated twice more to create three replicates of each of the 18 populations. Each of those populations was then sampled without replacement to produce, respectively, 2, 3, 4, 5, 6, and 7 sub-tests, each of (average) length K/t , where K = the number of items in the subscale (46 for Do, 34 for Sa, and 28 for Cm) and t = the number of sub-tests (2--7). (Selected population characteristics for these subscales are displayed in Table 1.) This resulted in a total of 324 individual multiple matrix samples, which were analyzed using the procedure given by Shoemaker (1973). Selected data output from Shoemaker's procedure ($\hat{\beta}$, the

estimate of the population mean; and $\hat{\sigma}^2$, the estimate of population variance) were used to compute errors of estimate of the mean ($EEM = \hat{\mu} - \mu$) and of the variance ($EEV = \hat{\sigma}^2 - \sigma^2$), where the parameters in each case were based on the populations of appropriate size.

Insert Table 1 about here

The relationship to EEM and EEV of N , t , and the skewness and kurtosis of each subscale/population combination was subsequently examined by performing an analysis of variance based on step-wise multiple linear regression (Kelly, Beggs, & McNeil, 1969; Kerlinger & Pedhazur, 1973). The procedure employed dummy variables with effect coding (Kerlinger & Pedhazur, 1973) to represent replications and subscale sources. The dummy variables were forced into the equation a priori, and the other variables mentioned, plus vectors representing the products of all possible combinations of the four major variables (representing interactions), were allowed to vie for inclusion in the regression equation on an equal basis under the default conditions of the SPSS REGRESSION program (Nie, Hull, Jenkins, Steinbrenner, & Bent, 1975).

Results and Discussion

The results for the regression equation for EEM are displayed in Table 2. The lack of a significant F ratio for the total equation (F_{overall}) indicates that the observed variations in EEM are not attributable to the variables under study. Thus it seems that neither the size of the population, the number of

sub-tests, the skewness or the kurtosis of the population distribution, nor any of their interactions, has a significant effect on the magnitude of the error of estimate of the mean.

Insert Table 2 about here.

Table 3, which pertains to EEV, however, yields a different conclusion. The overall F ratio was 24.799 ($d.f. = 15, 308; p < .001$). The population skewness had a significant effect on the error of estimate of the variance at the time of its entry into the regression equation ($F(3, 320) = 225.789, p < .001$), as did the interaction of skewness and kurtosis ($F(4, 319) = 79.183, p < .001$), the kurtosis itself ($F(5, 318) = 1.953, p < .05$), and the interaction of population size and kurtosis ($F(7, 316) = 3.249, p < .01$).

Insert Table 3 about here.

The significant contribution of the above-mentioned variables and the lack of significant contribution of the remaining variables suggested that a modification of the linear model underlying the analysis would be appropriate. Accordingly, the above-named four variables were used to generate a second regression equation which eliminated the other (non-significant) variables. This modification of the linear model results in the most conservative interpretation of the findings. As before, the replications data were forced into the regression equation first, then the four variables were allowed to vie for inclusion in the equation.

The results of the analysis based on the reduced linear model are in Table 4.

The effect of reducing the model by eliminating the non-significant variables is to increase the magnitude of MS (error); this in turn lessens the statistical significance of the effects of some of the variables. Under this model, the equation as a whole had an F of 60.879 (d.f. = 6, 317; p < .001), and the variables which retained significant effects, skewness and the skewness $\times$ kurtosis interaction, had F values of 81.263 (d.f. = 3, 320; p < .001) and 7.164 (d.f. = 4, 319; p < .001), respectively.

Insert Table 4 about here.

Study of the β weights in the final equation (see Table 4) yields the observation that the skewness and the skewness $\times$ kurtosis interaction are the two variables that are influential in determining the error of estimate of variance.

Interpreting the signs associated with the weights while recognizing the limitations imposed on interpretation by the relatively small number of scales and replications used in the study, we can tentatively conclude that increased positive skewness will decrease the error of estimate of the variance while increased negative skewness will increase the error of estimate of the variance. Furthermore, either increased positive skewness in conjunction with increased positive kurtosis (i.e., tendency toward leptokurtism) or increased negative skewness in conjunction with increased negative kurtosis (i.e., tendency toward platykurtism), will tend to increase the error of estimate of the variance, while increased positive skewness in conjunction with increased negative kurtosis or increased negative skewness in conjunction with increased positive kurtosis will tend to decrease the error of estimate of the variance.

The magnitude of the β weights associated with the skewness and the skewness * kurtosis interaction indicate that the two variables are approximately equally influential in determining the error of estimate of variance. Inasmuch as the skewness per se has influence on both variables, it could be expected that under certain circumstances (especially high negative skewness in a platykurtic distribution) it could greatly increase the error of estimate of variance.

Conclusions

The findings of this study with respect to errors of estimate of the mean indicate no effects due to either the number of sub-tests used, the size of the population, the skewness or kurtosis of the population, or any of the possible interactions of those variables. These findings are consistent with those of Barcikowski (1972, 1974), Moy and Barcikowski (1974), Scheetz and Forsyth (Note 1), and Shoemaker (1971). Thus the procedure for the estimation of the mean through MMS is apparently quite robust with respect to these population characteristics.

With respect to errors of estimate of the variance, this study contradicts earlier studies (generalized by Shoemaker, 1973; p. 22) by denying the relative importance of the number and length of subtests, whether with normal or skewed distributions. Although this finding could result from the limitations of this study described earlier, it could also reflect the multivariate analytic approach used here, in contrast to the essentially univariate approach taken in earlier studies.

The present study also indicates that the error of estimate of the variance is not independent of certain population characteristics. Forsyth (1976)

argues against Shoemaker's (1973) claim that the population skewness is influential in determining errors of estimate in MMS. The present study shows that skewness does exert considerable influence, both by itself, and through interaction with the kurtosis. Kurtosis per se, however, does not seem to be a significant factor.

Table 4 indicates that the variables under scrutiny in the reduced model account for only slightly over half the variance in the error of estimate of variance. Certainly sampling error would account for some of the remainder, as would inter-scale variability (the subscales used in this study differed markedly in terms of kurtosis, especially--see Table 1); the variance attributable to these variables could be reduced in future studies by increasing the numbers of replications and subscales used in the study, respectively. However, it is likely that even if this were done, a sizeable amount of unexplained variance would remain. This would suggest that perhaps other population characteristics might exist which influence the error of estimate of the variance. Barcikowski (1972, 1974; Moy & Barcikowsky, 1974) has identified one such possible characteristic in the item discrimination (i.e., the biserial correlation between item response and test score). Whether the effects of population skewness and the skewness $\times$ kurtosis interaction are as great as the effect of item discrimination can be determined by constructing a study similar to this one, and including both variables and associated interactions.

Reference Note

1. Scheetz, J. P., & Forsyth, R. A. Assigning items to sub-tests in multiple matrix sampling. Paper presented at the annual meeting of the American Educational Research Association, New York, April, 1977.

References

- Barcikowski, R. S. A Monte Carlo study of item sampling (versus traditional sampling) for norm construction. Journal of Educational Measurement, 1972, 9, 209-214.
- Barcikowski, R. S. The effects of item discrimination on the standard errors of estimate associated with item-examinee sampling procedures. Educational and Psychological Measurement, 1974, 34, 231-237.
- Cook, D. L., & Stufflebeam, D. L. Estimating test norms from variable size item and examinee samples. Educational and Psychological Measurement, 1967, 27, 601-610.
- Kelly, F. J., Beggs, D. L. & McNeil, K. A. Research design in the behavioral sciences: Multiple regression approach. Carbondale, Ill.: Southern Illinois University Press, 1969.
- Kerlinger, F. N., & Pedhazur, E. J. Multiple regression in behavioral research. New York: Holt, Rinehart & Winston, 1973.
- Lord, F. M. Estimating norms by item-sampling. Educational and Psychological Measurement, 1962, 22, 259-267.
- Lord, F. M., & Novick, M. R. Statistical theories of mental test scores. Reading, Mass.: Addison-Wesley, 1968.
- Moy, M. L. Y., & Barcikowski, R. S. Item sampling: Optimal number of people and items. The Journal of Experimental Education, 1974, 42(3), 46-52.
- Nie, N. H., Hull, C. H., Jenkins, J. G., Steinbrenner, K., & Bent, D. H. SPSS: Statistical package for the social sciences. (2nd ed.) New York: McGraw-Hill, 1975.

- Plumlee, L. B. Estimating means and standard deviations from partial data--an empirical check on Lord's item sampling technique. Educational and Psychological Measurement, 1964, 24, 623-630.
- Pugh, R. C. Empirical evidence on the application of Lord's sampling technique to Likert items. The Journal of Experimental Education, 1971, 39(3), 54-56.
- Scissons, E. H. Executive abilities: Implications for the counselling and training of middle and senior executives. Paper presented at the Adult Education Research Conference, Minneapolis, April, 1977. (ERIC Document Reproduction Service No. ED 141 509)
- Shoemaker, D. M. Allocation of items and examinees in estimating a norm distribution by item-sampling. Journal of Educational Measurement, 1970, 7, 123-128. (a)
- Shoemaker, D. M. Item-examinee sampling procedures and associated standard errors in estimating test parameters. Journal of Educational Measurement, 1970, 7, 255-262. (b)
- Shoemaker, D. M. Further results on the standard error of estimate associated with item-examinee sampling procedures. Journal of Educational Measurement, 1971 8, 215-220.
- Shoemaker, D. M. Standard errors of estimate in item-examinee sampling as a function of test reliability, variation in item difficulty indices and degree of skewness in the normative distribution. Educational and Psychological Measurement, 1972, 32, 705-714.
- Shoemaker, D. M. Principles and procedures of multiple matrix sampling. Cambridge, Mass.: Ballinger, 1973.

Shoemaker, D. M., & Okada, M. The communication skills program and concomitant spelling proficiency. Southwest Regional Laboratory for Educational Research and Development Technical Note, 1970. Cited by Shoemaker, D. M., Principles and procedures of multiple matrix sampling. Cambridge, Mass.: Ballinger, 1973.

Footnote

1. The author is indebted to Glen H. Wright for his assistance with the data analysis.

Table 1

Selected Characteristics of Populations

Scale	Population Size (N)	μ	σ^2	Skewness	Kurtosis
1 (Do)	75	32.960	31.471	-.933	.614
		33.440	19.304	-.985	1.353
		32.747	28.867	-.597	-.280
	100	33.210	30.006	-.832	.441
		33.400	21.232	-.910	1.365
		33.080	25.509	-.669	-.044
	150	33.160	27.061	-.731	.253
		33.627	21.376	-.709	.775
		33.087	24.912	-.618	-.048
	300	33.133	25.407	-.685	.280
		33.417	24.097	-.760	.560
		33.087	23.785	-.616	.164
	450	33.140	26.816	-.789	.474
		33.293	23.976	-.759	.558
		33.053	25.409	-.742	.414
	600	33.107	24.803	-.739	.437
2 (Sa)	75	22.747	9.543	-.205	-.065
		23.507	9.875	-.366	.316
		23.307	10.999	-.827	.310
	100	23.02	9.656	-.194	-.311
		23.56	10.067	-.150	.240
		23.420	10.468	-.602	.262
	150	22.980	9.805	-.109	-.338
		23.447	8.719	-.126	.419
		23.267	9.593	-.395	.098
	300	22.987	9.512	-.125	-.062
		23.250	9.312	-.147	-.027
		23.210	9.504	-.172	.191
	450	23.064	9.521	-.163	.072
		23.229	9.473	-.213	.120
		23.200	9.862	-.183	.136
	600	23.118	9.47	-.205	.090

continued...

Table 1 (continued)

Scale	Population Size (N)	μ	σ^2	Skewness	Kurtosis
3 (Cm)	75	26.16	2.136	-1.086	1.126
		26.173	2.470	-1.325	2.528
		26.373	2.534	-2.230	9.166
	100	26.280	2.022	-1.048	1.057
		26.100	2.212	-1.152	2.210
		26.270	2.623	-1.978	6.734
	150	26.193	2.144	-1.004	.989
		26.353	2.176	-1.329	2.445
		26.333	2.358	-1.929	6.396
	300	26.223	2.060	-.999	1.064
		26.243	2.185	-1.139	1.594
		26.117	2.552	-1.302	2.693
	450	26.193	2.579	-2.493	17.478
		26.164	2.810	-2.581	15.969
		26.178	2.387	-1.253	2.551
	600	26.19	2.688	-2.245	13.187
Average		27.530	12.431	-.882	2.450
S.D.		4.220	9.750	.677	4.461

Table 2
Analysis of Variance in Errors of Estimate of the Mean
Using Full Linear Model

$F_{\text{overall}} = .987 \quad (d.f. = 14, 309)^{\dagger}$

Source	MS_{reg}	df_{reg}	MS_{res}	df_{res}	$F_{\text{in}}^{\S}$	F_{fin}^{*}	R
Replications	4.097	2	3.887	321	1.054		.081
Number of Sub-tests	5.453	3	3.882	320	1.405	8.000 ³	.104
Population Size	3.617	4	3.883	319		3.775 ²	.117
Population Size × Number of Sub-tests	1.554	5	3.891	318		3.653 ²	.122
Number of Sub-tests × Skewness	.431	6	3.901	317		7.072 ³	.123
Skewness	3.066	7	3.904	316		7.331 ³	.133
Population Size × Number of Sub-tests × Skewness	1.062	8	3.913	315		3.589 ²	.136
Population Size × Skewness	3.020	9	3.916	314		3.316 ²	.145
Kurtosis	.390	10	3.927	313		6.684 ³	.146

[†] Overall F for the regression equation with all variables listed included.

[§] F values shown are those associated with the variable at the time of its entry into the equation.
 $F < 1.0$ omitted.

^{*} F value associated with the variable when all variables are included in the equation. $F < 1.0$ omitted;
 F for replications not included.

¹ $p < .05$

² $p < .01$

³ $p < .001$

Table 2 (continued)

Source	MS_{reg}	df_{reg}	MS_{res}	df_{res}	$F_{in}^{\S}$	F_{fin}^{*}	R
Skewness $\times$ Kurtosis	5.555	11	3.922	312	1.416	5.511 ³	.160
Number of Sub-tests $\times$ Kurtosis	4.232	12	3.921	311	1.079	5.405 ³	.171
Number of Sub-tests $\times$ Population Size $\times$ Skewness $\times$ Kurtosis	1.399	13	3.929	310		4.320 ³	.174
Population Size $\times$ Kurtosis	15.812	14	3.891	309	4.065	4.065 ³	.207

† Overall F for the regression equation with all variables listed included.

§ F values shown are those associated with the variable at the time of its entry into the equation.
 $F < 1.0$ omitted.

* F value associated with the variable when all variables are included in the equation. $F < 1.0$ omitted;
 F for replications not included.

¹ $p < .05$

² $p < .01$

³ $p < .001$

Table 3
Analysis of Variance in Errors of Estimate of Variance

Using Full Linear Model

$F_{\text{overall}} = 24.799^3$ (d.f. = 15, 308)[†]

Source	MS_{reg}	df_{reg}	MS_{res}	df_{res}	$F_{\text{in}}^{\S}$	F_{fin}^*	R
Replications	21.142	2	53.573	321			.050
Skewness	7114.252	3	31.508	320	225.789 ³	1.138	.644
Skewness × Kurtosis	2005.056	4	25.322	319	79.183 ³	4.886 ²	.729
Kurtosis	49.306	5	25.246	318	1.953 ¹	5.527 ²	.731
Skewness × Kurtosis × Population Size	27.850	6	25.238	317	1.103	2.951 ²	.732
Population Size × Kurtosis	81.433	7	25.060	316	3.249 ²	4.442 ³	.735
Population Size × Skewness	36.832	8	25.023	315	1.472		.737
Population Size × Number of Sub-tests	7.239	9	25.080	314			.737
Number of Sub-Tests × Skewness	11.630	10	25.123	313		1.055	.737

[†] Overall F for the regression equation with all variables listed included.

[§] F values shown are those associated with the variable at the time of its entry into the equation.
 $F < 1.0$ omitted.

^{*} F value associated with the variable when all variables are included in the equation. $F < 1.0$ omitted;
 F for replications not included.

¹ $p < .05$

² $p < .01$

³ $p < .001$

Table 3 (continued)

Source	MS_{reg}	df_{reg}	MS_{res}	df_{res}	$F_{in}^{\dagger}$	F_{fin}^*	R
Number of Sub-tests × Skewness × Kurtosis	13.013	11	25.161	312		1.197	.738
Number of Sub-tests	20.951	12	25.175	311			.739
Number of Sub-tests × Population Size × Skewness × Kurtosis	17.969	13	25.198	310			.740
Population Size × Number of Sub-tests × Skewness	2.399	14	25.272	309			.740
Population Size	.551	15	25.352	308			.740

† Overall F for the regression equation with all variables listed included.

§ F values shown are those associated with the variable at the time of its entry into the equation.
F < 1.0 omitted.

* F value associated with the variable when all variables are included in the equation. F < 1.0 omitted;
F for replications not included.

¹ p < .05

² p < .01

³ p < .001

Table 4
Analysis of Variance in Errors of Estimate of Variance
Using Reduced Linear Model

$F_{\text{overall}} = 60.879^1$ (d.f. = 6,317)[†]

Source	MS_{reg}	df_{reg}	MS_{res}	df_{res}	F_{fin}^*	$\underline{R}$	β
Replications	21.142	2	53.573	321		.050	.112 ²
Skewness	7114.252	3	31.508	320	81.263 ¹	.44	-1.138
Skewness × Kurtosis	2005.056	4	25.322	319	7.164 ¹	.729	1.262
Kurtosis	49.306	5	25.246	318		.731	.571
Population Size × Kurtosis	18.581	6	25.267	317		.732	.104

† Overall F for the regression equation with all variables listed included.

* F value associated with the variable when all variables are included in the equation. $F < 1.0$ omitted; F for replications not included.

¹ $p < .001$

² β weights for replications have been combined.